

Hate Speech's Rise on Twitter Is Unprecedented, Researchers Find



Elon Musk, who took over Twitter in October, has described himself as a “free speech absolutist.” Credit...Carina Johansen/NTB, via Reuters

By Sheera Frenkel and Kate Conger

Dec. 2, 2022

Before Elon Musk bought Twitter, slurs against Black Americans showed up on the social media service an average of 1,282 times a day. After the billionaire became Twitter’s owner, they jumped to 3,876 times a day.

Slurs against gay men appeared on Twitter 2,506 times a day on average before Mr. Musk took over. Afterward, their use rose to 3,964 times a day.

And antisemitic posts referring to Jews or Judaism soared more than 61 percent in the two weeks after Mr. Musk acquired the site.

These findings — from the Center for Countering Digital Hate, the Anti-Defamation League and other groups that study online platforms — provide the most comprehensive picture to date of how conversations on Twitter have changed since Mr. Musk completed his \$44 billion deal for the company in late October. While the numbers are relatively small, researchers said the increases were atypically high.

The shift in speech is just the tip of a set of changes on the service under Mr. Musk. Accounts that Twitter used to regularly remove — such as those that identify as part of the Islamic State, which were

banned after the U.S. government classified ISIS as a terror group — have come roaring back. Accounts associated with QAnon, a vast far-right conspiracy theory, have paid for and received verified status on Twitter, giving them a sheen of legitimacy.

These changes are alarming, researchers said, adding that they had never seen such a sharp increase in hate speech, problematic content and formerly banned accounts in such a short period on a mainstream social media platform.

“Elon Musk sent up the Bat Signal to every kind of racist, misogynist and homophobe that Twitter was open for business,” said Imran Ahmed, the chief executive of the Center for Countering Digital Hate. “They have reacted accordingly.”



Mr. Musk’s changes include allowing former President Donald J. Trump back on Twitter, which had barred him for tweets that could incite violence. Credit...Haiyun Jiang/The New York Times

Mr. Musk, who did not respond to a request for comment, has been vocal about being a “free speech absolutist” who believes in unfettered discussions online. He has moved swiftly to overhaul Twitter’s practices, allowing former President Donald J. Trump — who was barred for tweets that could incite violence — to return. Last week, Mr. Musk proposed a widespread amnesty for accounts that Twitter’s previous leadership had suspended. And on Tuesday, he ended enforcement of a policy against Covid misinformation.

But Mr. Musk has denied claims that hate speech has increased on Twitter under his watch. Last month, he tweeted a downward-trending graph that he said showed that “hate speech impressions”

had dropped by a third since he took over. He did not provide underlying numbers or details of how he was measuring hate speech.

Inside Elon Musk's Twitter

- **Licensing Music:** Twitter is said to have explored the licensing of music rights from three major labels before negotiations stalled after Elon Musk's takeover of the company last year.
- **Twitter's Finances:** Mr. Musk said that Twitter was recovering financially after seeing a 50 percent decline in ad revenue, making one of his first public disclosures about the state of the social media company since he purchased the company last year.
- **F.T.C. Investigation:** The Federal Trade Commission is intensifying an investigation into Twitter's data and privacy practices and is seeking testimony from Mr. Musk.
- **Job Cuts:** Twitter is said to have laid off at least 200 more workers. The social media platform now has a work force of fewer than 2,000 employees, down from 7,500 when Mr. Elon Musk took it over.

On Thursday, Mr. Musk said the account of Kanye West, which was restricted for a spell in October because of an antisemitic tweet, would be suspended indefinitely after the rapper, known as Ye, tweeted an image of a swastika inside the Star of David. On Friday, Mr. Musk said Twitter would publish "hate speech impressions" every week and agreed with a tweet that said hate speech spiked last week because of Ye's antisemitic posts.

Changes in Twitter's content not only have societal implications but also affect the company's bottom line. Advertisers, which provide about 90 percent of Twitter's revenue, have reduced their spending on the platform as they wait to see how it will fare under Mr. Musk. Some have said they are concerned that the quality of discussions on the platform will suffer.

On Wednesday, Twitter sought to reassure advertisers about its commitment to online safety. "Brand safety is only possible when human safety is the top priority," the company wrote in a blog post. "All of this remains true today."

The appeal to advertisers coincided with a meeting between Mr. Musk and Thierry Breton, the digital chief of the European Union, in which they discussed content moderation and regulation, according to an E.U. spokesman. Mr. Breton has pressed Mr. Musk to comply with the Digital Services Act, a European law that requires social platforms to reduce online harm or face fines and other penalties.



Thierry Breton, a European Union official, plans to visit Twitter for a “stress test” of its moderation. Credit...Johanna Geron/Reuters

Mr. Breton plans to visit Twitter’s San Francisco headquarters early next year to perform a “stress test” of its ability to moderate content and combat disinformation, the spokesman said.

On Twitter itself, researchers said the increase in hate speech, antisemitic posts and other troubling content had begun before Mr. Musk loosened the service’s content rules. That suggested that a further surge could be coming, they said.

If that happens, it’s unclear whether Mr. Musk will have policies in place to deal with problematic speech or, even if he does, whether Twitter has the employees to keep up with moderation. Mr. Musk laid off, fired or accepted the resignations of more than half the company’s staff last month, including those who worked to remove harassment, foreign interference and disinformation from the service. Yoel Roth, Twitter’s head of trust and safety, was among those who quit.

The Anti-Defamation League, which files regular reports of antisemitic tweets to Twitter and keeps track of which posts are removed, said the company had gone from taking action on 60 percent of the tweets it reported to only 30 percent.

“We have advised Musk that Twitter should not just keep the policies it has had in place for years, it should dedicate resources to those policies,” said Yael Eisenstat, a vice president at the Anti-Defamation League, who met with Mr. Musk last month. She said he did not appear interested in taking the advice of civil rights groups and other organizations.

“His actions to date show that he is not committed to a transparent process where he incorporates the best practices we have learned from civil society groups,” Ms. Eisenstat said. “Instead he has emboldened racists, homophobes and antisemites.”

The lack of action extends to new accounts affiliated with terror groups and others that Twitter previously banned. In the first 12 days after Mr. Musk assumed control, 450 accounts associated with ISIS were created, up 69 percent from the previous 12 days, according to the Institute for Strategic Dialogue, a think tank that studies online platforms.

Other social media companies are also increasingly concerned about how content is being moderated on Twitter.

When Meta, which owns Facebook and Instagram, found accounts associated with Russian and Chinese state-backed influence campaigns on its platforms last month, it tried to alert Twitter, said two members of Meta’s security team, who asked not to be named because they were not authorized to speak publicly. The two companies often communicated on these issues, since foreign influence campaigns typically linked fake accounts on Facebook to Twitter.

But this time was different. The emails to their counterparts at Twitter bounced or went unanswered, the Meta employees said, in a sign that those workers may have been fired.